Eigenvalues and eigenvectors

Let $A \in \mathbb{R}^{n \times n}$. If $0 \neq v \in \mathbb{C}^n$ and $\lambda \in \mathbb{C}$ satisfy

$$Av = \lambda v$$

then λ is called **eigenvalue**, and v is called **eigenvector**.

Given a matrix, we want to approximate its eigenvalues and eigenvectors. Some applications:

- Structural engineering (natural frequency, heartquakes)
- Electromagnetics (resonance cavity)
- Google's Pagerank algorithm
- ...

The characteristic polynomial

The eigenvalues of a matrix are the roots of **the characteristic polynomial**

$$p(\lambda) := \det(\lambda I - A) = 0$$

However, computing the roots of a polynomial is a very ill-conditioned problem! We cannot use this approach to compute the eigenvalues.

Eigenvalues and eigenvectors

Algorithms that compute the eigenvalues/eigenvectors of a matrix are divided into two categories:

1. Methods that compute all the eigenvalues/eigenvectors at once.
2. Methods that compute only a few (possibly one) eigenvalues/eigenvectors.

The methods are also different whether the matrix is symmetric or not. In this lesson we will discuss methods of type 2.

Diagonalizable matrices

Definition

We say that a matrix $A \in \mathbb{C}^{n \times n}$ is diagonalizable if there exists a non singular matrix U and a diagonal matrix D such that $U^{-1}AU = D$.

The diagonal element of D are the eignevalue of A and the column u_i of U is an eigenvector of A relative to the eigenvalue $D_{i,i}$.

Since a scalar multiple of an eigenvector is still an eigenvector, we can choose U such that $\|u_i\|_2 = 1$ for $i = 1, \ldots, n$.

Finally, we observe that if A is diagonalizable, since U is non singular, then the vectors $\{u_1, \ldots, u_n\}$ form a basis of $\mathbb{C}^n$.

From now on, we assume that the eigenvalues are numbered in decreasing order (in module), i.e.

$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n|$$

Eigenvalues/eigenvectors of a symmetric matrix

Theorem

All the eigenvalues of a real symmetric matrix are **real**. Moreover, there exists a basis of eigenvectors $u_1, \dots, u_n$, i.e.

$$Au_i = \lambda_i u_i$$

that are orthonormal, i.e.

$$(u_i, u_j) = \delta_{ij}$$

The power method

We want to approximate the eigenvalue of A that is largest in module.

$v_0 =$ some vector with $\|v_0\| = 1$.
for k $= 1, 2, \ldots$

$w = Av_{k-1}$ apply A
$v_k = w / \|w\|$ normalize
$\mu_k = (v_k)^H A v_k$ Reyleigh quotient

end

Theorem

Let $A \in \mathbb{C}^{n \times n}$ be a diagonalizable matrix. Assume $|\lambda_1| > |\lambda_2|$ and $v_0 = \sum_{i=1}^{n} \alpha_i u_i$, with $\alpha_1 \neq 0$. Then there exists $C > 0$, independent of k, such that

$$\|\widetilde{v}_k - u_1\|_2 \leq C \left| \frac{\lambda_2}{\lambda_1} \right|^k, \qquad \text{where } \widetilde{v}_k = \frac{\|A^k v_0\|}{\alpha_1 \lambda_1^k} v_k.$$

Proof

We expand v_0 on the eigenvector basis $\{u_1, \ldots, u_n\}$ choosen s.t. $\|u_i\| = 1$ for $i = 1, \ldots, n$:

$$v_0 = \sum_{i=1}^{n} \alpha_i u_i, \qquad \text{with } \alpha_1 \neq 0$$

It holds

$$A^k v_0 = \sum_{i=1}^{n} \alpha_i \lambda_i^k u_i \qquad \text{and} \qquad v_k = \frac{A^k v_0}{\|A^k v_0\|}$$

Hence, we can write

$$\widetilde{v}_k = \frac{A^k v_0}{\alpha_1 \lambda_1^k} = u_1 + \sum_{i=2}^{n} \frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1}\right)^k u_i$$

At this point, it holds

$$\|\widetilde{v}_k - u_1\|_2 = \left\|\sum_{i=2}^{n} \frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1}\right)^k u_i\right\|_2 \leq \sum_{i=2}^{n} \left\|\frac{\alpha_i}{\alpha_1} \left(\frac{\lambda_i}{\lambda_1}\right)^k u_i\right\|_2 = \sum_{i=2}^{n} \left|\frac{\alpha_i}{\alpha_1}\right| \left|\frac{\lambda_i}{\lambda_1}\right|^k$$

So, we obtain

$$\|\widetilde{v}_k - u_1\|_2 \leq \sum_{i=2}^{n} \left|\frac{\alpha_i}{\alpha_1}\right| \left|\frac{\lambda_i}{\lambda_1}\right|^k \leq (n-1) \cdot \max_{i=2,\dots,n} \left(\left|\frac{\alpha_i}{\alpha_1}\right|\right) \left|\frac{\lambda_2}{\lambda_1}\right|^k = C \left|\frac{\lambda_2}{\lambda_1}\right|^k,$$

where we have defined $C = (n-1) \cdot \max_{i=2,\dots,n} \left(\left|\frac{\alpha_i}{\alpha_1}\right|\right)$. Since C does not depend on k, this concludes the proof.

The previous theorem implies that the sequence $\{\widetilde{v}_k\}$ converges to the eigenvector u_1. Since $\widetilde{v}_k$ is a scalar multiple of v_k, they have the same direction and this direction converges to the direction of u_1. As a result, for k that goes to $+\infty$ the vector v_k tends to have the same direction of u_1,. Thus v_k tends to be an eigenvector relaltive to λ_1.

Remark

if $|\lambda_2| \ll |\lambda_1|$ the convergence will be fast. On the other hand, if $\lambda_2 \approx \lambda_1$ the convergence will be slow.